BC

#7 OIPE

RAW SEQUENCE LISTING DATE: 09/24/2001
PATENT APPLICATION: **US/09/700,708** TIME: 15:34:05

Input Set : **A:\P20294.sL1.ST25.txt**
Output Set: **N:\CRF3\09242001\I700708.raw**

<110> APPLICANT: ITAI, Akiko
ITAI, Reiko
TOMIOKA, Nobuo
<120> TITLE OF INVENTION: Method For Predicting Functions of Protein
<130> FILE REFERENCE: P20294
<140> CURRENT APPLICATION NUMBER: 09/700,708
<141> CURRENT FILING DATE: 2000-11-24
<150> PRIOR APPLICATION NUMBER: PCT/JP98/02302
<151> PRIOR FILING DATE: 1998-05-26
<160> NUMBER OF SEQ ID NOS: 6
<170> SOFTWARE: PatentIn version 3.0
<210> SEQ ID NO: 1
<211> LENGTH: 159
<212> TYPE: PRT
<213> ORGANISM: Escherichia coli
<400> SEQUENCE: 1

```
Met Ile Ser Leu Ile Ala Ala Leu Ala Val Asp Arg Val Ile Gly Met
1               5                  10                  15
Glu Asn Ala Met Pro Trp Asn Leu Pro Ala Asp Leu Ala Trp Phe Lys
            20                  25                  30
Arg Asn Thr Leu Asp Lys Pro Val Ile Met Gly Arg His Thr Trp Glu
        35                  40                  45
Ser Ile Gly Arg Pro Leu Pro Gly Arg Lys Asn Ile Ile Leu Ser Ser
    50                  55                  60
Gln Pro Gly Thr Asp Asp Arg Val Thr Trp Val Lys Ser Val Asp Glu
65                  70                  75                  80
Ala Ile Ala Ala Cys Gly Asp Val Pro Glu Ile Met Val Ile Gly Gly
                85                  90                  95
Gly Arg Val Tyr Glu Gln Phe Leu Pro Lys Ala Gln Lys Leu Tyr Leu
            100                 105                 110
Thr His Ile Asp Ala Glu Val Glu Gly Asp Thr His Phe Pro Asp Tyr
        115                 120                 125
Glu Pro Asp Asp Trp Glu Ser Val Phe Ser Glu Phe His Asp Ala Asp
    130                 135                 140
Ala Gln Asn Ser His Ser Tyr Cys Phe Lys Ile Leu Glu Arg Arg
145                 150                 155
```

<210> SEQ ID NO: 2
<211> LENGTH: 223
<212> TYPE: PRT
<213> ORGANISM: Bovine
<400> SEQUENCE: 2

```
Ile Val Gly Gly Tyr Thr Cys Gly Ala Asn Thr Val Pro Tyr Gln Asx
1               5                  10                  15
Ser Leu Asn Ser Gly Tyr His Phe Cys Gly Gly Ser Leu Ile Asn Ser
            20                  25                  30
Gln Trp Val Val Ser Ala Ala His Cys Tyr Lys Ser Gly Ile Gln Val
        35                  40                  45
```

RAW SEQUENCE LISTING DATE: 09/24/2001
PATENT APPLICATION: US/09/700,708 TIME: 15:34:05

Input Set : A:\P20294.sL1.ST25.txt
Output Set: N:\CRF3\09242001\I700708.raw

```
74 Arg Leu Gly Glu Asp Asn Ile Asn Val Asx Glu Gly Asn Glu Gln Phe
75     50                  55                  60
77 Ile Ser Ala Ser Lys Ser Ile Val His Pro Ser Tyr Asn Ser Asn Thr
78 65                  70                  75                  80
80 Leu Asn Asn Asp Ile Met Leu Ile Lys Leu Lys Ser Ala Ala Ser Leu
81                 85                  90                  95
83 Asn Ser Arg Val Ala Ser Ile Ser Leu Pro Thr Ser Cys Ala Ser Ala
84             100                 105                 110
86 Gly Thr Gln Cys Leu Ile Ser Gly Trp Gly Met Thr Lys Ser Ser Gly
87         115                 120                 125
89 Thr Ser Tyr Pro Asp Asx Leu Lys Cys Leu Lys Ala Pro Ile Leu Ser
90     130                 135                 140
92 Asp Ser Ser Cys Lys Ser Ala Tyr Pro Gly Gln Ile Thr Ser Asn Met
93 145                 150                 155                 160
95 Phe Cys Ala Gly Tyr Leu Glu Gly Gly Lys Asp Ser Cys Gln Gly Asp
96                 165                 170                 175
98 Cys Gly Gly Pro Val Val Cys Ser Gly Lys Leu Gln Gly Ile Val Ser
99             180                 185                 190
101 Trp Gly Ser Gly Cys Ala Gln Lys Asn Lys Pro Gly Val Tyr Thr Lys
102         195                 200                 205
104 Val Cys Asn Tyr Val Ser Trp Ile Lys Gln Thr Ile Ala Ser Asn
105     210                 215                 220
107 <210> SEQ ID NO: 3
108 <211> LENGTH: 124
109 <212> TYPE: PRT
110 <213> ORGANISM: Bovine
112 <400> SEQUENCE: 3
114 Lys Glu Thr Ala Ala Ala Lys Phe Glu Arg Gln His Met Asp Ser Ser
115 1               5                   10                  15
117 Thr Ser Ala Ala Ser Ser Ser Asn Tyr Cys Asn Gln Met Met Lys Ser
118             20                  25                  30
120 Arg Asn Leu Thr Lys Asp Arg Cys Lys Pro Val Asn Thr Phe Val His
121         35                  40                  45
123 Glu Ser Leu Ala Asp Val Gln Ala Val Cys Ser Gln Lys Asn Val Ala
124     50                  55                  60
126 Cys Lys Asn Gly Gln Thr Asn Cys Tyr Gln Ser Tyr Ser Thr Met Ser
127 65                  70                  75                  80
129 Ile Thr Asp Cys Arg Glu Thr Gly Ser Ser Lys Tyr Pro Asn Cys Ala
130                 85                  90                  95
132 Tyr Lys Thr Thr Gln Ala Asn Lys His Ile Ile Val Ala Cys Glu Gly
133             100                 105                 110
135 Asn Pro Tyr Val Pro Val His Phe Asp Ala Ser Val
136         115                 120
138 <210> SEQ ID NO: 4
139 <211> LENGTH: 153
140 <212> TYPE: PRT
141 <213> ORGANISM: Whale
143 <400> SEQUENCE: 4
145 Val Leu Ser Glu Gly Glu Trp Gln Leu Val Leu His Val Trp Ala Lys
```

RAW SEQUENCE LISTING DATE: 09/24/2001
PATENT APPLICATION: **US/09/700,708** TIME: 15:34:05

Input Set : **A:\P20294.sL1.ST25.txt**
Output Set: **N:\CRF3\09242001\I700708.raw**

```
146 1               5                   10                  15
148 Val Glu Ala Asp Val Ala Gly His Gly Gln Asp Ile Leu Ile Arg Leu
149             20                  25                  30
151 Phe Lys Ser His Pro Glu Thr Leu Glu Lys Phe Asp Arg Phe Lys His
152         35                  40                  45
154 Leu Lys Thr Glu Ala Glu Met Lys Ala Ser Glu Asp Leu Lys Lys His
155     50                  55                  60
157 Gly Val Thr Val Leu Thr Ala Leu Gly Ala Ile Leu Lys Lys Lys Gly
158 65                  70                  75                  80
160 His His Glu Ala Glu Leu Lys Pro Leu Ala Gln Ser His Ala Thr Lys
161                 85                  90                  95
163 His Lys Ile Pro Ile Lys Tyr Leu Glu Phe Ile Ser Glu Ala Ile Ile
164             100                 105                 110
166 His Val Leu His Ser Arg His Pro Gly Asp Phe Gly Ala Asp Ala Gln
167         115                 120                 125
169 Gly Ala Met Asn Lys Ala Leu Glu Leu Phe Arg Lys Asp Ile Ala Ala
170     130                 135                 140
172 Lys Tyr Lys Glu Leu Gly Tyr Gln Gly
173 145                 150
175 <210> SEQ ID NO: 5
176 <211> LENGTH: 186
177 <212> TYPE: PRT
178 <213> ORGANISM: Human
180 <400> SEQUENCE: 5
182 Val Gly Ser Leu Asn Cys Ile Val Ala Val Ser Gln Asn Met Gly Ile
183 1               5                   10                  15
185 Gly Lys Asn Gly Asp Leu Pro Trp Pro Pro Leu Arg Asn Glu Phe Arg
186             20                  25                  30
188 Tyr Phe Gln Arg Met Thr Thr Thr Ser Ser Val Glu Gly Lys Gln Asn
189         35                  40                  45
191 Leu Val Ile Met Gly Lys Lys Thr Trp Phe Ser Ile Pro Glu Lys Asn
192     50                  55                  60
194 Arg Pro Leu Lys Gly Arg Ile Asn Leu Val Leu Ser Arg Glu Leu Lys
195 65                  70                  75                  80
197 Glu Pro Pro Gln Gly Ala His Phe Leu Ser Arg Ser Leu Asp Asp Ala
198                 85                  90                  95
200 Leu Lys Leu Thr Glu Gln Pro Glu Leu Ala Asn Lys Val Asp Met Val
201             100                 105                 110
203 Trp Ile Val Gly Gly Ser Ser Val Tyr Lys Glu Ala Met Asn His Pro
204         115                 120                 125
206 Gly His Leu Lys Leu Phe Val Thr Arg Ile Met Gln Asp Phe Glu Ser
207     130                 135                 140
209 Asp Thr Phe Phe Pro Glu Ile Asp Leu Glu Lys Tyr Lys Leu Leu Pro
210 145                 150                 155                 160
212 Glu Tyr Pro Gly Val Leu Ser Asp Val Gln Glu Glu Lys Gly Ile Lys
213                 165                 170                 175
215 Tyr Lys Phe Glu Val Tyr Glu Lys Asn Asp
216             180                 185
218 <210> SEQ ID NO: 6
```

RAW SEQUENCE LISTING DATE: 09/24/2001
PATENT APPLICATION: **US/09/700,708** TIME: 15:34:05

Input Set : **A:\P20294.sL1.ST25.txt**
Output Set: **N:\CRF3\09242001\I700708.raw**

```
219 <211> LENGTH: 186
220 <212> TYPE: PRT
221 <213> ORGANISM: Escherichia coli
223 <400> SEQUENCE: 6
225 Val Gly Ser Leu Ser Leu Ile Ala Ala Leu Ala Gln Asn Met Gly Ile
226 1               5                   10                  15
228 Gly Lys Asn Gly Asp Leu Pro Trp Pro Pro Leu Pro Ala Asp Leu Ala
229             20                  25                  30
231 Trp Phe Lys Arg Asn Thr Leu Asp Lys Ser Val Glu Gly Lys Gln Asn
232         35                  40                  45
234 Leu Val Ile Met Gly Arg His Thr Trp Glu Ser Ile Gly Arg Pro Leu
235     50                  55                  60
237 Pro Gly Arg Lys Gly Arg Ile Asn Leu Val Leu Ser Arg Glu Leu Lys
238 65                  70                  75                  80
240 Glu Pro Pro Gln Gly Ala His Phe Leu Ser Arg Ser Leu Asp Asp Ala
241                 85                  90                  95
243 Leu Lys Leu Thr Glu Gln Pro Glu Leu Ala Asn Lys Val Asp Met Val
244             100                 105                 110
246 Met Val Ile Gly Gly Gly Ser Val Tyr Lys Glu Ala Met Asn His Pro
247         115                 120                 125
249 Gly His Leu Lys Leu Tyr Leu Thr His Ile Met Gln Asp Phe Glu Ser
250     130                 135                 140
252 Asp Thr Phe Phe Pro Glu Ile Asp Leu Glu Lys Tyr Lys Leu Leu Pro
253 145                 150                 155                 160
255 Glu Tyr Pro Gly Val Leu Ser Asp Val Gln Glu Glu Lys Gly Ile Lys
256                 165                 170                 175
258 Tyr Lys Phe Glu Val Tyr Glu Lys Asn Asp
259             180                 185
```

VERIFICATION SUMMARY DATE: 09/24/2001
PATENT APPLICATION: **US/09/700,708** TIME: 15:34:06

Input Set : **A:\P20294.sL1.ST25.txt**
Output Set: **N:\CRF3\09242001\I700708.raw**